(19) JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: 11122296 A

(43) Date of publication of application: 30.04.99

(51) Int. Cl H04L 12/56

(21) Application number: 09277449

(22) Date of filing: 09.10.97

(71) Applicant: CHOKOSOKU NETWORK COMPUTER GIJUTSU KENKYUSHO:KK

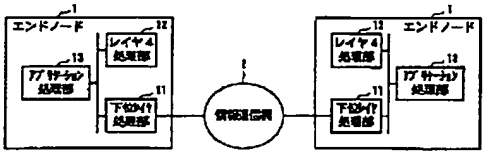
(72) Inventor: ATSUMI YUKIO

(54) BAND CONTROL METHOD

(57) Abstract:

PROBLEM TO BE SOLVED: To realize detailed service based on a priority set as to the use of a communication band.

SOLUTION: A layer 4 processing section 12 executes a communication protocol to set a communication band used for packet transmission usually based on the smaller value of a reception notice window value denoting reception capability of an opposite node and a congestion window value obtd. by estimating a transfer capability of a network. On the detection of a packet missing even, the congestion window value is not reduced immediately but stored as it is and only in the case that a succeeding packet missing state is deteriorated, the congestion window value is reduced to reduce the communication band.



Scope of Claims

Claim 1: A bandwidth control method, in a communication protocol which sets a communication bandwidth used in packet transmission based on the smaller value among a reception advertised window value indicating the reception capacity of the remote node, and a congestion window value which estimates the transmission capacity of the network; characterized in that

upon detection of a packet loss event, the congestion window value is not immediately reduced, but is maintained without change; and,

only when the subsequent packet loss situation worsens is the congestion window value reduced and the communication bandwidth reduced.

Claim 2: The bandwidth control method described in Claim 1, characterized in that

a judgment is made as to whether preferential communication, in which bandwidth is allocated preferentially according to the user level or application type, is necessary;

in cases where preferential communication is unnecessary, upon detection of a packet loss event, the congestion window value is immediately reduced;

in cases where preferential communication is necessary, upon detection of a packet loss event, the congestion window value is not immediately reduced but is maintained without change; and,

only when the subsequent packet loss situation worsens is the congestion window value reduced and the communication bandwidth reduced.

Claim 3: The bandwidth control method described in Claim 1 or in Claim 2, characterized in that

a packet loss amount is computed for each prescribed monitoring interval;

a first packet loss amount in an arbitrary monitoring interval and a second packet loss amount in the immediately succeeding monitoring interval are compared; and,

when the second packet loss amount is larger than the first packet loss amount, the packet loss situation is judged to be worsening.

Claim 4: The bandwidth control method described in Claim 3, characterized in that a monitoring interval is taken to be the interval from the time of detection of a packet loss event until the time of delivery confirmation notification from the receiving side for all transmitted packets.

Claim 5: A bandwidth control method, in a communication protocol in which throughput indicating the network load situation is measured for each measurement interval of a prescribed length, and packet transmission amounts are adjusted according to changes therein; characterized in that

upon detection of a lowering of throughput, the packet transmission amount is not immediately reduced but is maintained without change; and,

only when the subsequent throughput lowering worsens is the packet transmission amount reduced and the communication bandwidth reduced.

Claim 6: The bandwidth control method described in Claim 5, characterized in that:

a judgment is made as to whether preferential communication, in which bandwidth is allocated preferentially according to the user level or application type, is necessary;

in cases where preferential communication is unnecessary, upon detection of a lowering of throughput, the packet transmission amount is immediately reduced;

in cases where preferential communication is necessary, upon detection of a lowering of throughput, the packet transmission amount is not immediately reduced but is maintained without change; and,

only when the throughput lowering further worsens is the packet transmission amount reduced and the communication bandwidth reduced.

Claim 7: The bandwidth control method described in Claim 5 or in Claim 6, characterized in that:

a first throughput change amount is computed from the difference between the throughput of an arbitrary first measurement interval and the throughput of the immediately succeeding second measurement interval;

when the computed first throughput change amount is greater than or equal to a prescribed threshold value, the throughput is judged to be declining;

a second throughput change amount is computed from the difference between the throughput of the second measurement interval and the throughput of the immediately succeeding third measurement interval; and,

when the computed first and second throughput change amounts are equal to or greater than a prescribed threshold value, the lowering of throughput is judged to be further worsening.

Detailed Description of the Invention

[0001]

Technical Field of the Invention

This invention relates to a bandwidth control method, and in particular relates to a bandwidth control method for cases in which the bandwidth of a protocol layer 4 is controlled at a node executing communication protocol processing.

[0002]

Prior Art

At present, principal services on the Internet and on intranets take the form of best-effort services, and bandwidth is provided at each connection according to the network situation, which can change from moment to moment. In TCP,

which is a representative communication protocol, changes in the network situation are detected based on detection of packet loss events and similar, and the transmission amount is adjusted.

[0003]

For this reason, in addition to the reception advertised window value (adwnd) which is a parameter indicating the reception capacity of the remote node, a congestion window value (cwnd) which is a parameter indicating an estimated value of the network transmission capacity is used for transmission in the range min(adwnd,cwnd). Here min(A,B) indicates the smaller of the values of A and B.

[0004]

Bandwidth control to adjust transmission amounts is realized primarily through cwnd, and in addition to the above-described parameters, there is also a slow start threshold (ssthresh). Packet transmission is performed based on these parameters; but there are two phases, called the slow start phase and the congestion avoidance phase, with different control methods.

[0005]

A summary of the method is as follows. As the units of each parameter for bandwidth control, in the TCP implementation, number of bytes is used; but in order to facilitate the following explanation, the number of packets is employed. First, after setting the connection, the bandwidth

control parameters are initially set to ssthresh=adwnd and cwnd=1.

[0006]

The slow start phase is then entered, one data packet (DT packet) is transmitted, and reception of an acknowledgement response packet (ACK packet) is awaited. When an ACK packet is received within a fixed amount of time, cwnd is incremented by one, and then two packets are transmitted; subsequently, each time an ACK packet is received the cwnd which is the amount that can be transmitted next is increased by the number of DT packets acknowledged, until the threshold ssthresh is reached.

[0007]

Hence the possible transmission amount cwnd increases in the manner 1, 2, 4, 8, ... When cwnd reaches ssthresh, the congestion avoidance phase is entered. In the congestion avoidance phase, each time an ACK packet is received cwnd is increased by the amount 1/cwnd, so that compared with the slow start phase, the rate of increase is gradual.

[0008]

In a representative TCP implementation, a DT packet loss is judged to have occurred when the time monitor times out without having received a confirmation response, or when a constant number (normally 3) or more duplicate ACKs are received. Duplicate ACKs are received when a plurality of continuous ACK packets having the same receive sequence number are received.

[0009]

When DT packet loss is judged to have occurred, the bandwidth control parameters are adjusted, and ssthresh is set to min(cwnd,adwnd)/2. When duplicate ACKs are detected, the new cwnd is set equal to the old cwnd/2, and upon a time out, the new cwnd is set equal to 1 (see for example W.R. Stevens, TCP Illustrated Vol. 1, Chapter 21, Addison Wesley, 1994).

[0010]

On the other hand, a function for selective response (selective ACK, or SACK) is described as a TCP option in RFC2018, a 1996 Internet document, with the aim of speeding error recovery. On the receiving side, SACK information, for all packets for which reception is discontinuous due to packet loss, is added to ACK packets during intervals when reception is possible, to provide specific notification. On the transmitting side, this information is used to resend lost packets.

[0011]

In relation to bandwidth control (congestion control), existing functions are to be preserved in RFC2018, so that when packet loss is detected using SACK information, parameter adjustment is performed similarly to when the duplicate ACKs described above are detected. In this way, when duplicate ACKs, response time out, or packet loss detection through SACK information occur, control is performed to uniformly reduce

the transmission amount, regardless of the user level or application type using the connection.

[0012]

Problems to be Solved by the Invention

However, in such a bandwidth control method of the prior art, and particularly in the TCP implemented by various current hosts and terminal devices, during congested network conditions such as result in packet loss events, a best-effort service is provided which uniformly reduces communication bandwidth for each connection regardless of communication request conditions of applications and users, and consequently there is the problem that, during congested network conditions, bandwidth cannot be maintained for communication connections with high priority, such as for example important or urgent [communication].

[0013]

That is, together with the diversification and sophistication of information communication services, there is an increasing need in best-effort services to perform prioritization for use of communication bandwidth as one condition for requesting so-called QoS (Quality of Service), to provide more finely-tuned services. For example, in an intranet which presupposes use within a company, it is possible to perform prioritization of communication bandwidth allocation according to business type and job description, in order to secure throughput of important communications and

urgent communications and to shorten response times. This is not possible using current TCP bandwidth control.

[0014]

This invention is intended to resolve such problems, and has as an object the provision of a bandwidth control method enabling the realization of more finely-tuned services based on prioritization of the utilization of communication bandwidth.

[0015]

Means to Solve the Problems

In order to achieve this object, among the bandwidth control methods of this invention, the invention of claim 1 is [a method] which, in a communication protocol which sets the communication bandwidth used in packet transmission based on the smaller value among a reception advertisement window value indicating the reception capacity of the remote node and a congestion window value which estimates the transmission capacity of the network, upon detection of a packet loss event, the congestion window value is not immediately reduced but is maintained without change, and only in cases where subsequently the packet loss situation worsens is the congestion window reduced and the communication bandwidth reduced. Hence upon detection of a packet loss event, the communication bandwidth is not immediately reduced but is preserved in the current state, and only in cases where the

subsequent packet loss situation worsens is the communication bandwidth reduced.

[0016]

The invention of claim 2 is the invention of claim 1, wherein a judgment is made as to whether preferential communication, in which bandwidth is allocated preferentially according to the user level or application type, is necessary; in cases where preferential communication is unnecessary, upon detection of a packet loss event, the congestion window value is immediately reduced; in cases where preferential communication is necessary, upon detection of a packet loss event, the congestion window value is not immediately reduced but is maintained without change; and, only when the subsequent packet loss situation worsens is the congestion window value reduced and the communication bandwidth reduced. Hence upon detection of a packet loss event, the communication bandwidth is immediately reduced for connections for which preferential communication is unnecessary, the communication bandwidth released by this means can be used to preserve the communication bandwidth of connections requiring preferential communication in the current state, and only when subsequently the packet loss situation worsens is the communication bandwidth of connections requiring preferential communication decreased.

[0017]

The invention of claim 3 is the invention of claim 1 or claim 2, wherein a packet loss amount is computed for each prescribed monitoring interval; a first packet loss amount in an arbitrary monitoring interval and a second packet loss amount in the immediately succeeding monitoring interval are compared; and, when the second packet loss amount is larger than the first packet loss amount, the packet loss situation is judged to be worsening. The invention of claim 4 is the invention of claim 3, wherein a monitoring interval is taken to be the interval from the time of detection of a packet loss event until the time of delivery confirmation notification from the receiving side for all transmitted packets.

[0018]

In the invention of claim 5, in a communication protocol in which throughput indicating the network load situation is measured for each measurement interval of a prescribed length, and packet transmission amounts are adjusted according to changes therein, upon detection of a lowering of throughput, the packet transmission amount is not immediately reduced but is maintained without change; and, only when the subsequent throughput lowering further worsens is the packet transmission amount reduced and the communication bandwidth reduced. Hence upon detection of throughput lowering, the communication bandwidth is not immediately reduced but is preserved unchanged, and only when subsequent throughput lowering further worsens is the communication bandwidth reduced.

[0019]

The invention of claim 6 is the invention of claim 5, wherein a judgment is made as to whether preferential communication, in which bandwidth is allocated preferentially according to the user level or application type, is necessary; in cases where preferential communication is unnecessary, upon detection of a lowering of throughput, the packet transmission amount is immediately reduced; in cases where preferential communication is necessary, upon detection of a lowering of throughput, the packet transmission amount is not immediately reduced but is maintained without change; and, only when the throughput lowering further worsens is the packet transmission amount reduced and the communication bandwidth reduced. Hence when throughput lowering is detected, the communication bandwidth is immediately reduced for connections for which preferential communication is unnecessary, the communication bandwidth resulting therefrom is used to preserve the communication bandwidth of connections requiring preferential communication in the current state, and only when subsequent throughput lowering further worsens is the communication bandwidth reduced for connections requiring preferential communication.

[0020]

The invention of claim 7 is the invention of claim 5 or claim 6, wherein a first throughput change amount is computed from the difference between the throughput of an arbitrary

first measurement interval and the throughput of the immediately succeeding second measurement interval; when the computed first throughput change amount is greater than or equal to a prescribed threshold value, the throughput is judged to be declining; a second throughput change amount is computed from the difference between the throughput of the second measurement interval and the throughput of the immediately succeeding third measurement interval; and, when the computed first and second throughput change amounts are equal to or greater than a prescribed threshold value, the lowering of throughput is judged to be further worsening.

[0021]

Aspects of the Invention

Next, this invention is explained referring to the drawings. Fig. 1 is a block diagram of a communication system which is a first aspect of this invention. In this figure, the end node 1 is connected to the data communication network 2, and communicates with the remote end node 1; the data communication network 2 is a network comprising communication circuits and relay nodes.

[0022]

The end node 1 comprises a lower layer processing portion 11, a layer 4 processing portion 12, and an application processing portion 13. The lower layer processing portion 11 performs the processing of the protocol layer 3 and below, that is, the processing of layer 1 including electrical

coordination with the communication network, the processing of layer 2 such as frame composition and decomposition, and the processing of layer 3 (here taken to be the IP) such as routing.

[0023]

The layer 4 processing portion 12 performs the processing of layer 4 (here taken to be TCP) to set and release connections and to transmit and receive data based on flow control. In this case, as parameters for bandwidth control for protocol layer 4 (TCP), in addition to the reception advertised window value (adwnd), congestion window value (cwnd) and slow start threshold (ssthresh), a function to count the number of lost packets, variables loss1, loss2 to store the number of lost packets, and a variable hi_chk to store the maximum unconfirmed transmission sequence number transmitted at the time of detection of packet loss, are also used.

[0024]

The number of packets which can be transmitted is min(cwnd,adwnd), and this number varies as each parameter varies during communication. Here min(A,B) indicates the smaller of the values of A and B. The upper layer has a preferential communication flag which can be turned on and off to select the method of bandwidth control in TCP.

[0025]

In a first aspect of this invention, when a packet loss event is detected in a connection for which the preferential communication flag is turned on, the bandwidth control parameters ssthresh and cwnd are not immediately reduced but are maintained, and only in cases where the subsequent loss situation worsens are the bandwidth control parameters ssthresh and cwnd reduced. The interval from the time of detection of a packet loss event until delivery is confirmed for all packets which have been transmitted and for which delivery has not been confirmed is taken to be the monitoring interval, and by comparing the total number of lost packets for two neighboring monitoring intervals, it is judged whether the loss situation after the detection of packet loss event is worsening or not.

[0026]

Next, the operation of data transmission, centered on bandwidth control, is explained as the operation of the first aspect of this invention, referring to Figs. 2 through 4. Fig. 2 is a sequence explanation drawing which indicates in summary the entirety of the communication sequence; Fig. 3 is a flowchart showing in summary the entirety of bandwidth control; and Fig. 4 is a state transition table which indicates bandwidth control processing during preferential communication. The following is an explanation for the example of a case in which a packet is transferred from the server side (transmission side) in response to a request from the client side (receiving side).

[0027]

First, the layer 4 processing portion 12 (TCP) on the client side exchanges connection setting requests and connection setting responses with the layer 4 processing portion 12 (TCP) on the server side, based on a communication initiation request from the application processing portion 13 (upper layer AP), and by this means makes connection settings. At this time, when there is a need to use the SACK option in the TCP connection, the SACK indication is appended to a connection setting request packet, and negotiation is performed by the transmitting side and the receiving side.

[0028]

The layer 4 processing portion 12 (TCP) on the client side exchanges connection release requests and connection release responses with the layer 4 processing portion 12 (TCP) on the server side, based on a communication termination request from the application processing portion 13 (upper AP), and by this means releases the connection.

[0029]

As indicated in Fig. 2, after completion of connection settings, the server side enters a normal (N) state in which data transmission is possible (step 31 in Fig. 3). The upper AP on the server side judges whether to perform preferential bandwidth control for the established TCP connection based on the user ID or requesting application ID as notified by the upper AP on the client side.

[0030]

When the ID of the notification indicates a connection requiring preferential communication, the preferential communication flag is turned on, and when a connection not requiring preferential communication is indicated, the flag is turned off. Hence in the later normal (N) state of data transfer, when a TCP DT packet loss is detected by the layer 4 processing portion 12 on the server side, whether preferential communication is required for this connection is judged according to whether the preferential communication flag is turned on or off.

[0031]

As shown in Fig. 3, in the normal (N) state of data transfer (step 31), when a TCP DT packet loss is detected by the server-side layer 4 processing portion 12, the preferential communication flag is checked (step 32). If the preferential communication flag is off, the connection is judged to be a connection for which normal bandwidth control is specified, similar to the prior art, and the bandwidth control parameters ssthresh and cwnd are immediately reduced (step 33).

[0032]

On the other hand, if the preferential communication flag is on, the connection is judged to be a connection for which preferential communication is necessary, and the following bandwidth control processing is executed. A packet loss event

can be apprehended from duplicate ACKs, from SACK (selective acknowledgement) information, and from transmission acknowledgement timeouts. The number of lost packets L is 1 in the case of duplicate ACKs and a transmission acknowledgement timeout, but in the case of SACK information, a new loss number L is ascertained from this information.

[0033]

In the normal (N) state, upon reception of duplicate ACKs or upon SACK reception, a monitoring interval is started, and at this time the maximum unacknowledged transmission sequence number for transmitted packets is set in the variable hi_chk, the number of lost packets L is set in the variable loss1, and there is a transition to the monitoring 1 (M1) state (step 34) without reduction of the cwnd or ssthresh values. In the N state, a monitoring interval is started even on delivery acknowledgement timeout; at this time the maximum unacknowledged transmission sequence number for transmitted packets is set in the variable hi_chk, the variable loss1 is set to 1, and there is a transition to the M1 state (step 34) without reduction of the values of cwnd or ssthresh.

[0034]

In the M1 state, processing upon ACK receipt is as follows. When the reception sequence number for the received ACK packet (ACKseq) is such that hi_chk > ACKseq, that is, when delivery has not been confirmed for any transmitted and

unconfirmed packets, loss1 is not changed, the M1 state is unchanged, and the monitoring interval is continued.

[0035]

On the other hand, when hi_chk ≤ ACKseq, that is, when delivery is confirmed for all transmitted and unconfirmed packets, at this time the maximum transmission sequence number for transmitted and unconfirmed packets is set in hi_chk, the variable loss2 is set to 0, and there is a transition to the monitoring 2 (M2) state (step 35). In this way, the first monitoring interval is ended, and the next monitoring interval is begun.

[0036]

In the M1 state, when duplicate ACKs are received or when a SACK is received, the following processing is performed. When the reception sequence number of the received ACK packet (ACKseq) is such that hi_chk > ACKseq, that is, when there has not yet been delivery confirmation for all transmitted but unconfirmed packets, if a new packet loss is detected the number of losses L is added to loss1, the M1 state is unchanged, and the monitoring interval is continued.

[0037]

On the other hand, when hi_chk ≤ ACKseq, that is, when delivery is confirmed for all transmitted and unconfirmed packets, at that time the maximum transmission sequence number for transmitted and unconfirmed packets is set in hi_chk,if

there is a new packet loss detected the loss number L is set in loss2, and there is a transition to the M2 state. In this way, the first monitoring interval is ended and the next monitoring interval is begun. In the M1 state, when there is a delivery confirmation timeout, loss1 is incremented by 1, the M1 state is unchanged, and the monitoring interval is continued.

[0038]

In the M2 state, processing upon ACK receipt is as follows. When the reception sequence number for the received ACK packet (ACKseq) is such that hi_chk > ACKseq, that is, when delivery has not yet been confirmed for all transmitted but unconfirmed packets, there is no change to loss2, the M2 state is unchanged, and the monitoring interval is continued.

[0039]

On the other hand, when hi_chk ≤ ACKseq, that is, when delivery is confirmed for all transmitted and unconfirmed packets, the two succeeding monitoring intervals are judged to be ended, and the values of loss1 and loss2 are compared (step 36). Here if loss2=0, no packet losses have occurred in the succeeding monitoring interval, and so the congestion situation is judged to be alleviated, and there is a transition to the normal (N) state (step 31) without a reduction of the bandwidth control parameters ssthresh and cwnd.

[0040]

If loss1 ≥ loss2, then the total number of lost packets has decreased in the succeeding monitoring interval, and so the congestion situation is considered to be improving; hence the bandwidth control parameters ssthresh and cwnd are not decreased, but a new monitoring interval is begun in order to further observe the situation. To this end, the value of loss1 is set to that of loss2, loss2 is set to 0, the maximum transmission sequence number for transmitted but unconfirmed packets is set to hi_chk, the M2 state (step 35) is re-entered, and the monitoring interval is continued.

[0041]

If loss1 < loss2, the total number of lost packets in the succeeding monitoring interval is increasing, and so the network congestion situation is considered to be worsening; hence after reducing the bandwidth control parameters ssthresh and cwnd (step 37), a transition to the N state (step 31) is made.

[0042]

In the M2 state, processing when duplicate ACKs are received or when a SACK is received is as follows. When the reception sequence number for a received ACK packet (ACKseq) is such that hi_chk > ACKseq, that is, when delivery has not been confirmed for all transmitted but unconfirmed packets, if a new packet loss is detected, the number of losses L is added to loss2, the M2 state is unchanged, and the monitoring interval is continued.

[0043]

When hi_chk ≤ ACKseq, that is, when delivery is confirmed for all transmitted and unconfirmed packets, the two succeeding monitoring intervals are judged to be ended, and the values of loss1 and loss2 are compared (step 38). Here if loss1 ≥ loss2, the total number of packet losses in the succeeding monitoring interval is decreasing, and so it is considered that the congestion situation is improving, and the bandwidth control parameters ssthresh and cwnd are not reduced.

[0044]

Also, in order to further observe the situation, loss1 is set to the value of loss2, hi_chk is set to the maximum transmission sequence number for transmitted and unconfirmed packets, and when a new packet loss is detected, loss2 is set to the number of losses L, the M2 state is resumed (step 35), and a new monitoring interval is begun.

[0045]

If loss1 < loss2, then the total number of lost packets in the succeeding monitoring interval has increased, and so it is considered that the network congestion situation is worsening; hence after reducing the bandwidth control parameters ssthresh and cwnd (step 39), there is a transition to the N state (step 31).

[0046]

Further, when there is a delivery confirmation timeout in the M2 state, loss2 is incremented by one, the M2 state remains unchanged, and the monitoring interval is continued. Reduction of the bandwidth control parameters ssthresh and cwnd in steps 36 and 38 is at this time similar to that of the prior art, with ssthresh = min(cwnd,adwnd)/2, and the new cwnd = the old cwnd/2.

[0047]

Thus in the first aspect of this invention, when a packet loss event is detected, the bandwidth control parameters ssthresh and cwnd are not immediately reduced but are maintained, and only when the subsequent loss situation worsens are the bandwidth control parameters ssthresh and cwnd reduced. Hence compared with a case in which the bandwidth is reduced uniformly in response to detection of packet loss events, as in conventional bandwidth control methods such as shown in Fig. 5(a), the bandwidth can be preserved for a connection for which communication is preferred, due to importance, urgency, or some other reason, as shown in Fig. 5(b), and so a finely-tuned service can be realized based on the assignment of preferences regarding utilization of communication bandwidth.

[0048]

Further, the interval from the time of detection of a packet loss event until the time of delivery confirmation for all packets which have been transmitted but delivery of which

is unconfirmed is taken to be a monitoring interval, and the total numbers of lost packets loss1, loss2 during an arbitrary monitoring interval and the succeeding monitoring interval are compared, so that it is possible to accurately judge whether, after detection of a packet loss event, the loss situation is worsening or not.

[0049]

Based on a user ID or requesting application ID provided by the receiving side (the client side), a judgment is made on the transmitting side (the server side) as to whether preferential bandwidth control is to be performed for the connection; hence unlimited preferential communication for numerous receiving [nodes] can easily be suppressed, and in addition bandwidth can immediately be reduced in response to detection of packet loss events, similarly to the prior art in which preferential communication is not employed, so that bandwidth can be secured for preferential communications.

[0050]

Next, a second aspect of this invention is explained. Here the network congestion situation is apprehended not in terms of packet loss events, but in terms of changes in throughput, and an explanation is given for the case of a communication protocol which adjusts transmission amounts (for example, a TCP version called TCP-Vegas). The approach is similar to that described above; here a brief explanation is given.

[0051]

First, throughput is calculated for each measurement interval of a prescribed length. When computing throughput, first the time required from the start of packet transmission in an arbitrary measurement interval until notification of delivery confirmation from the receiving side for all transmitted packets transmitted during the measurement interval is measured. The transmitted packet amount during this measurement interval is then divided by the required time to compute the throughput.

[0052]

Next, the throughput computed for an arbitrary measurement interval is taken to be P1 and the throughput computed for the immediately succeeding measurement interval is taken to be P2, and the amount of change in throughput D between the two measurement intervals is computed by taking the difference of the two, that is, D = P1-P2. Here a prescribed threshold value a is established, and is used to judge whether throughput is declining or not.

[0053]

Operation for the case in which the preferential communication flag is turned on is here explained. If $D < a$, then the throughput is judged to be satisfactory, and the amount of packet transmission is increased. If $D \geq a$, throughput is judged to be declining, the current transmission

amount value is maintained, and the situation in the next interval is observed.

[0054]

Next, the P2 value is saved as P1, and the throughput for the next interval is determined as P2. The throughput change amount D is again computed and is compared with the threshold a, and if again $D \geq a$, it is judged that the throughput lowering has worsened further, and the current packet transmission amount is reduced.

[0055]

On the other hand, operation when the preferential communication flag is off is as follows. The throughput change amount D, computed as described above, and the threshold a are compared, and if $D < a$, throughput is judged to be satisfactory, and the transmission amount is increased. If $D \geq a$, throughput is declining, and the network load is considered to be increasing, so that the current transmission amount value is immediately reduced.

[0056]

Hence in a communication protocol which apprehends [network conditions] in terms of throughput changes and adjusts transmission amounts, when a decline in throughput is detected, the packet transmission amount is not reduced immediately but is maintained; only when the decline in throughput subsequently worsens is the packet transmission

amount reduced. Consequently, compared with cases in which the bandwidth is reduced uniformly in response to detection of packet loss events, such as in the prior art, bandwidth can be maintained for a connection for which communication is preferred, due to importance, urgency, or some other reason, so that a finely-tuned service can be realized based on the assignment of preferences regarding utilization of communication bandwidth.

[0057]

Advantageous Results of the Invention

As explained above, when in this invention a network congestion situation is detected, the transmission amount is not immediately decreased for connections performing preferential communication but is maintained, and only when the subsequent congestion situation further worsens is the transmission bandwidth of connections performing preferential communication reduced. Hence compared with cases in which bandwidth is reduced uniformly in response to detection of packet loss events as in the prior art, bandwidth can be maintained for a connection for which communication is preferred, due to importance, urgency, or some other reason, and in addition the diversification of QoS (Quality of Service) in a best-effort type service can be accommodated, to realize a finely-tuned service based on the assignment of preferences regarding utilization of communication bandwidth. For example, in an intranet which assumes use within a company,

it is possible to perform prioritization of communication bandwidth allocation according to business type and job description, in order to secure throughput of important communications and urgent communications and to shorten response times, enabling realization of an aggressive best-effort service.

[0058]

Further, a judgment is made as to whether preferential communication, in which bandwidth is allocated preferentially according to the user level or application type, is necessary; in cases where preferential communication is unnecessary, upon detection of congestion, the packet transmission amount is immediately reduced; in cases where preferential communication is necessary, upon detection of congestion, the packet transmission amount is not immediately reduced but is maintained without change; and, only when the subsequent congestion situation worsens is the packet transmission amount reduced and the communication bandwidth reduced. Hence even when congestion occurs in the network and the usable communication bandwidth is reduced, the communication bandwidth of connections not requiring preferential communication and for which [bandwidth] is immediately reduced in response to the event of congestion is employed to maintain in its current state the communication bandwidth of connections requiring preferential communication. Whether preferential communication, in which bandwidth is assigned preferentially according to the user level or application type,

is necessary or not can be judged by the transmitting side, so that unlimited preferential communication for numerous receiving [nodes] can easily be suppressed, and in addition bandwidth can immediately be reduced in response to detection of packet loss events, similarly to the prior art in which preferential communication is not employed, so that bandwidth can be secured for preferential communications.

[0059]

Further, a packet loss amount is computed for each prescribed monitoring interval; a first packet loss amount in an arbitrary monitoring interval and a second packet loss amount in the immediately succeeding monitoring interval are compared; and, when the second packet loss amount is larger than the first packet loss amount, the packet loss situation is considered to be worsening. Also, the interval from the time of detection of a packet loss event until the time of delivery confirmation notification from the receiving side for all transmitted packets, so that in a communication protocol which detects the network congestion situation through packet loss events and which adjusts the packet transmission amount, it is possible to accurately judge whether, after detection of a packet loss event, the loss situation is worsening or not.

[0060]

Further, a first throughput change amount is computed from the difference between the throughput of an arbitrary first measurement interval and the throughput of the

immediately succeeding second measurement interval; when the computed first throughput change amount is greater than or equal to a prescribed threshold value, the throughput is judged to be declining; a second throughput change amount is computed from the difference between the throughput of the second measurement interval and the throughput of the immediately succeeding third measurement interval; and, when the computed first and second throughput change amounts are equal to or greater than a prescribed threshold value, the lowering of throughput is judged to be further worsening. Thus in a communication protocol in which the network congestion situation is detected from changes in throughput and the packet transmission amount is adjusted, declines in throughput, and subsequent further worsening thereof, can be accurately judged.

Brief Description of the Drawings

Fig. 1 is a block diagram of a data communication system which is one aspect of this invention.

Fig. 2 is a sequence explanation drawing which shows in summary the entirety of the communication sequence.

Fig. 3 is a flowchart which shows in summary the entirety of bandwidth control.

Fig. 4 is a state transition table which shows bandwidth control processing during preferential communication.

Fig. 5 is an explanatory drawing which shows the concept of operations in this invention.

Explanation of Symbols

1 end node

2 data communication network

11 lower layer processing portion

12 layer 4 processing portion

13 application processing portion

Fig. 1

1 end node

11 lower layer processing portion

12 layer 4 processing portion

13 application processing portion

2 data communication network

Fig. 4

a··· State
e··· Transition factor
f··· ACK receipt

b··· Normal (N)

g··· *Conventional operation (slow start, transmission operation of congestion avoidance phase)
*Remains in state N

C··· Monitoring 1 (M1)

h··· (1) When hi_chk > ACKseq
*Remains in state M1
(2) When hi_chk ≤ ACKseq
*Set hi_chk
*loss2 set to 0
*Change to state M2

D··· Monitoring 2 (M2)

i··· (1) When hi_chk > ACKseq
*Remains in state M2
(2) When hi_chk ≤ ACKseq
(a) When loss2=0
*Change to state N
(b) When loss1 ≥ loss2
*Loss1 set to loss2
*Set hi_chk
*loss2 set to 0
*Remains in state M2

j... Duplicate ACK receipt, or SACK receipt	k... *Set hi_chk *Set loss1 to L *Change to state M1	l... (1) When hi_chk > ACKseq *Set loss1 to loss1+1 *Remains in state M1 (2) When hi_chk ≤ ACKseq *Set hi_chk *Set loss2 to L *Change to state M2	m... (1) When hi_chk > ACKseq *Set loss2 to loss2+L *Remain in state M2 (2) When hi_chk ≤ ACKseq (a) When loss1 ≥ loss2 *Set loss1 to loss2 *Set hi_chk *Set loss2 to L *Remain in state M2 (c) When loss1 < loss2 *Reduce ssthresh, cwnd *Change to state N
n... Delivery confirmation timeout	o... *Set hi_chk *Set loss1 to 1 *Change to state M1	p... *Set loss1 to loss1+1 *Remain in state M1	q... *Set loss2 to loss2+1 *Remain in state M2	

r... Notes

Set hi_chk: Set the variable hi_chk to the maximum transmission sequence number for transmitted but unconfirmed [packets]

L: Number of new lost packets (in the case of duplicate ACKs this is 1; in the case of a SACKit is the newly determined number of lost packets)

Fig. 5

(a) Conventional bandwidth control method

a... Transmitting side

c... Network

b... Receiving side

d... Connection A

e ... Connection B

f ... Connection C

g ... Loss detection

h ... Congestion

a' ... Transmitting side

c' ... Network

b' ... Receiving side

d' ... Connection A

e' ... Connection B

f' ... Connection C

(b) Bandwidth control method of this invention

A ... Transmitting side

C ... Network

B ... Receiving side

D ... Connection A

E ... Connection B

F ... Connection C

G ... Loss detection

H ... Congestion

A' ... Transmitting side

C' ... Network

B' ... Receiving side

D'... Connection A

E'... Connection B

F'... Connection C

G... Connection for which this invention is implemented

K... Note: The connection thickness denotes the size of the bandwidth.

Fig. 2

a... Server side

b... Client side

C... Upper AP

D... Network

E... Upper AP

F... Communication initiation request

G... Connection setting request

H... Connection setting response

I... User ID or requesting application ID

J... Preferential communication?

K... Turn on preferential communication flag

L... High-priority transmission control

M... Change in requesting application

N... Preferential communication?

O... Turn off preferential communication flag

P... Conventional transmission control

Q... Communication termination request

R... Connection release request

S... Connection release response

Fig. 3

31 Normal (N)

a... Loss event

32 Preferential communication flag

33 Reduce ssthresh, cwnd

b... During M1 period (loss1 counting)

34 Monitoring 1 (M1)

C... M1 period ends

d... During M2 period (loss2 counting)

35 Monitoring 2 (M2)

e... M2 period ends (ACK received)

36 Compare loss1 and loss2

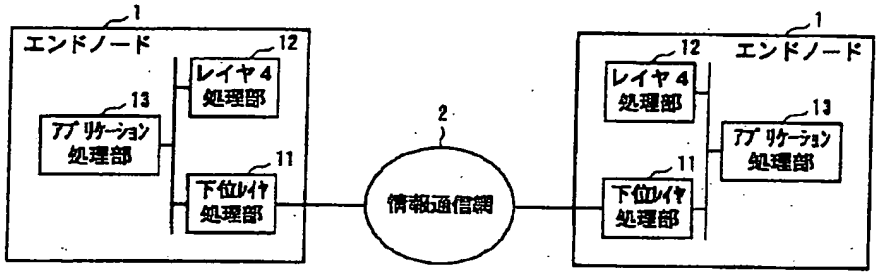
37 Reduce ssthresh, cwnd

f... M2 period ends (duplicate ACKs received, SACK received)

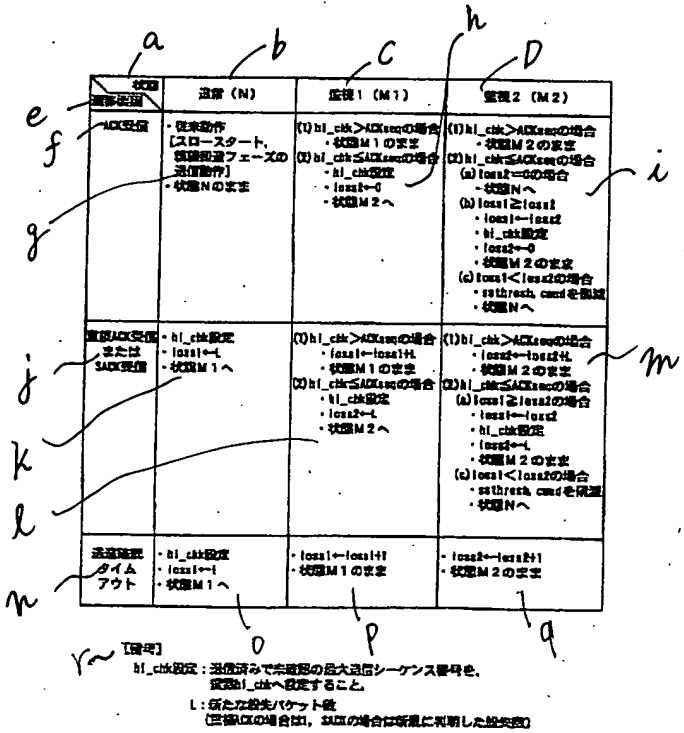
38 Compare loss1 and loss2

39 Reduce ssthresh, cwnd

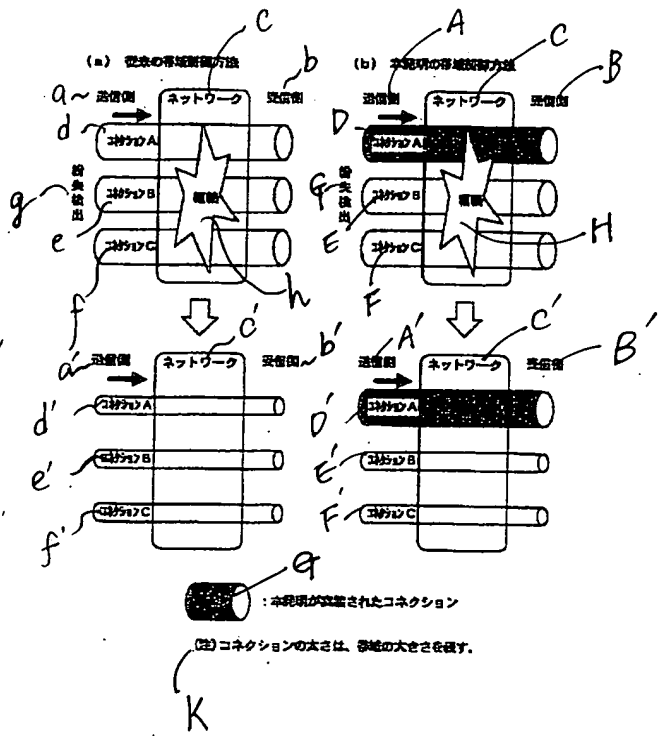
【図1】 FIG. 1



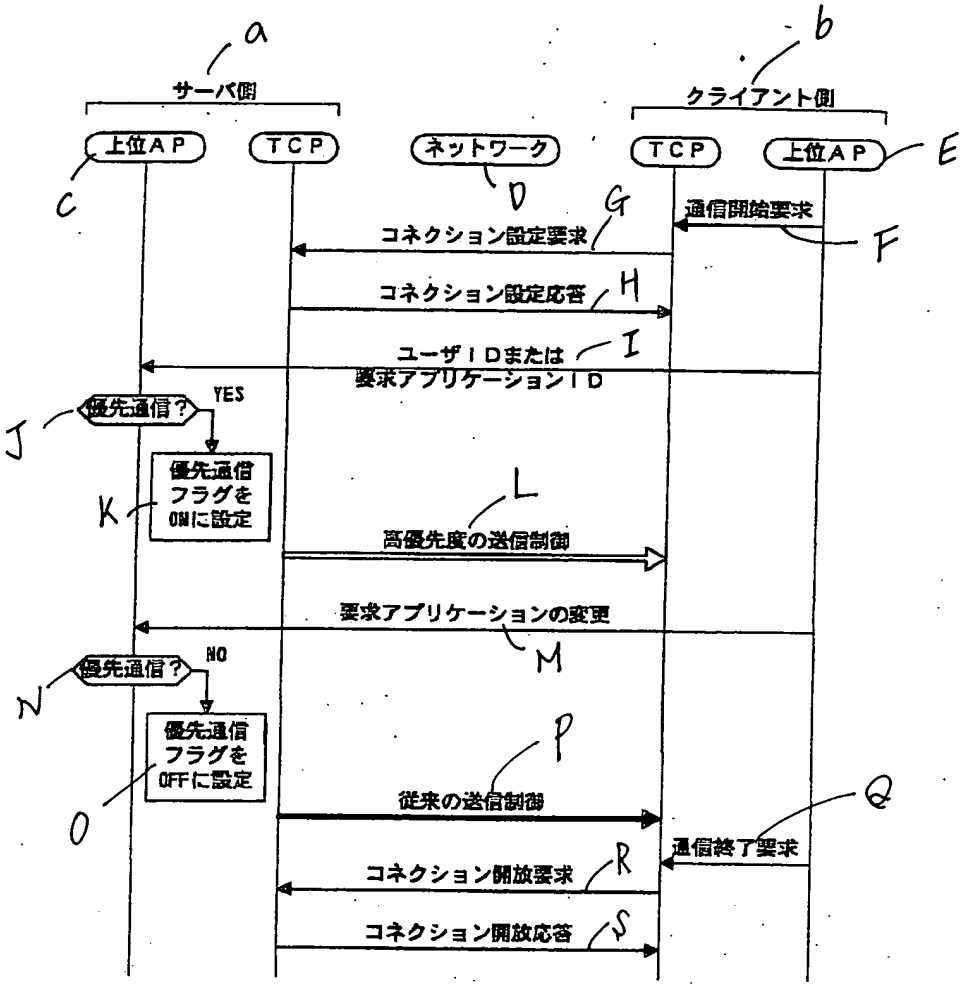
【図4】 FIG. 4



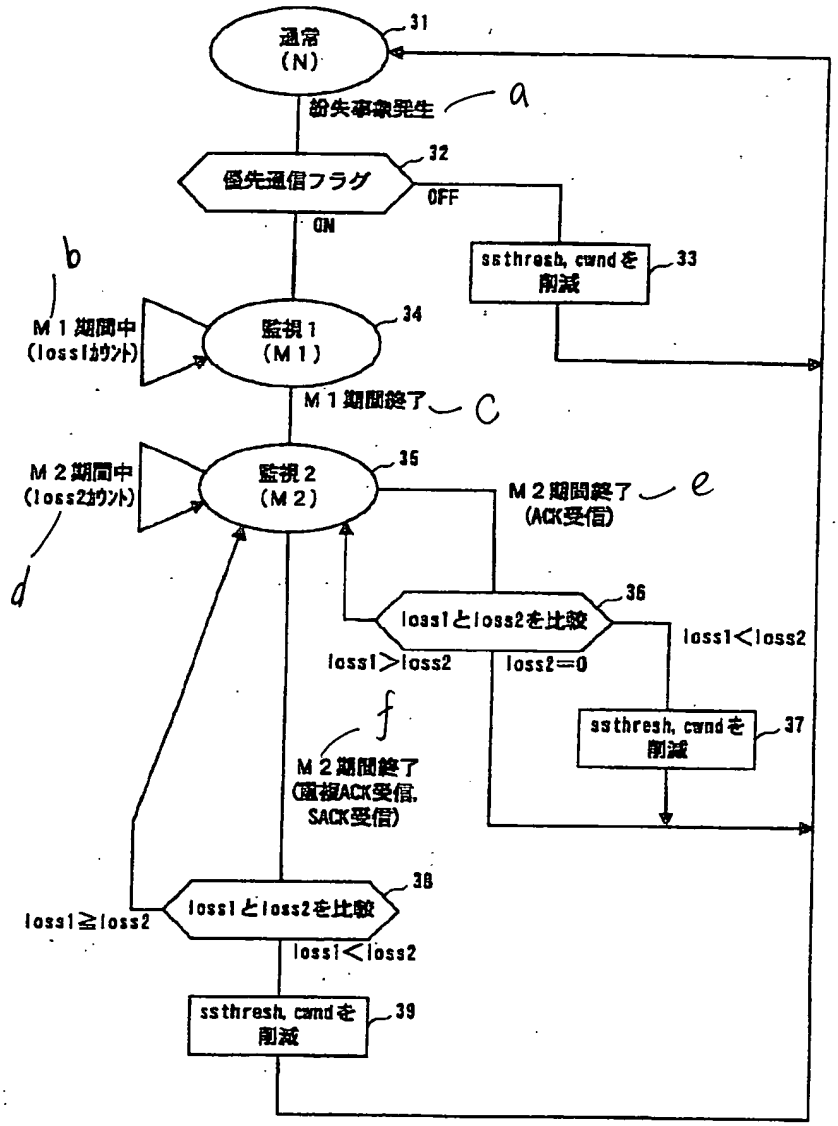
【図5】 FIG. 5



[図2] FIG. 2


a
b
サーバ側
クライアント側
上位AP
TCP
ネットワーク
TCP
上位AP
C
D
E
G
通信開始要求
F
コネクション設定要求
コネクション設定応答
H
ユーザIDまたは
要求アプリケーションID
I
優先通信？
YES
J
優先通信
フラグを
ONに設定
K
L
高優先度の送信制御
要求アプリケーションの変更
M
優先通信？
NO
N
優先通信
フラグを
OFFに設定
O
P
従来の送信制御
Q
通信終了要求
コネクション開放要求
R
コネクション開放応答
S

【図3】 FIG. 3


通常
(N)
31
a
32
優先通信フラグ
OFF
ON
ssthresh, cwndを
削減
33
b
M1期間中
(loss1カウント)
監視1
(M1)
34
M1期間終了
c
M2期間中
(loss2カウント)
d
監視2
(M2)
35
M2期間終了
(ACK受信)
e
36
loss1とloss2を比較
loss1<loss2
loss1>loss2
loss2=0
f
M2期間終了
(重複ACK受信,
SACK受信)
ssthresh, cwndを
削減
37
38
loss1とloss2を比較
loss1≧loss2
loss1<loss2
ssthresh, cwndを
削減
39

(19) JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **11122296 A**

(43) Date of publication of application: **30.04.99**

(51) Int. Cl **H04L 12/56**

(21) Application number: **09277449**

(22) Date of filing: **09.10.97**

(71) Applicant: **CHOKOSOKU NETWORK COMPUTER GIJUTSU KENKYUSHO:KK**

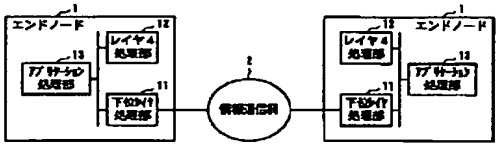
(72) Inventor: **ATSUMI YUKIO**

(54) **BAND CONTROL METHOD**

(57) Abstract:

PROBLEM TO BE SOLVED: To realize detailed service based on a priority set as to the use of a communication band.

SOLUTION: A layer 4 processing section 12 executes a communication protocol to set a communication band used for packet transmission usually based on the smaller value of a reception notice window value denoting reception capability of an opposite node and a congestion window value obtd. by estimating a transfer capability of a network. On the detection of a packet missing even, the congestion window value is not reduced immediately but stored as it is and only in the case that a succeeding packet missing state is deteriorated, the congestion window value is reduced to reduce the communication band.



(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（A）　(11)特許出願公開番号

特開平11－122296

(43)公開日　平成11年(1999) 4月30日

(51)Int.Cl.[6]	識別記号	FI	
H04L 12/56		H04L 11/20	102C

審査請求 有　請求項の数7　OL　(全 10 頁)

(21)出願番号　特願平9－277449

(22)出願日　平成9年(1997)10月9日

(71)出願人　394025577
株式会社超高速ネットワーク・コンピュータ技術研究所
東京都港区虎ノ門五丁目2番6号

(72)発明者　渥美　幸雄
東京都港区虎ノ門5丁目2番6号　株式会社超高速ネットワーク・コンピュータ技術研究所内

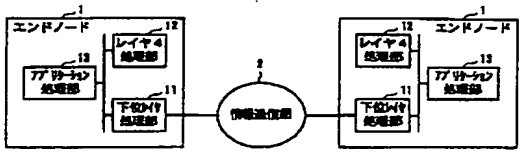
(74)代理人　弁理士　山川　政樹

(54)【発明の名称】　帯域制御方法

(57)【要約】

【課題】　通信帯域の利用についての優先度付けに基づいたきめ細かいサービスを実現する。

【解決手段】　レイヤ４処理部１２では、通常、相手ノードの受信能力を示す受信告知ウィンドウ値と、ネットワークの転送能力を推定した輻輳ウィンドウ値とのいずれか小さい値に基づいて、パケット送信に用いる通信帯域を設定する通信プロトコルを実行する。ここで、パケット紛失事象検出時には、輻輳ウィンドウ値を直ちには削減せずにそのまま保持し、その後のパケット紛失状況が悪化した場合にのみ輻輳ウィンドウ値を削減して通信帯域を削減する。



1

【特許請求の範囲】

【請求項１】　相手ノードの受信能力を示す受信告知ウィンドウ値と、ネットワークの転送能力を推定した輻輳ウィンドウ値とのいずれか小さい値に基づいて、パケット送信に用いる通信帯域を設定する通信プロトコルにおいて、

パケット紛失事象検出時には、輻輳ウィンドウ値を直ちには削減せずにそのまま保持し、

その後のパケット紛失状況が悪化した場合にのみ輻輳ウィンドウ値を削減して通信帯域を削減することを特徴とする帯域制御方法。

【請求項２】　請求項１記載の帯域制御方法において、

ユーザレベルあるいはアプリケーション種別に応じて帯域を優先的に割当てる優先通信が必要か否か判断し、

優先通信が不要な場合には、パケット紛失事象検出時に輻輳ウィンドウ値を直ちに削減し、

優先通信が必要な場合には、パケット紛失事象検出時に輻輳ウィンドウ値を直ちには削減せずにそのまま保持し、

その後のパケット紛失状況が悪化した場合にのみ輻輳ウィンドウ値を削減して通信帯域を削減することを特徴とする帯域制御方法。

【請求項３】　請求項１または２記載の帯域制御方法において、

所定の監視区間ごとにパケット紛失量を算出し、

任意の監視区間での第１のパケット紛失量と、その直後の監視区間での第２のパケット紛失量とを比較し、

第２のパケット紛失量が第１のパケット紛失量より大きい場合には、パケット紛失状況が悪化していると判断することを特徴とする帯域制御方法。

【請求項４】　請求項３記載の帯域制御方法において、

パケット紛失事象検出時点から送信済みのパケットすべてについて受信側から送達確認通知された時点までの区間を監視区間とすることを特徴とする帯域制御方法。

【請求項５】　ネットワークの負荷状況を示すスループットを所定長の測定区間ごとに測定し、その変化に応じてパケット送信量を調整する通信プロトコルにおいて、

スループットの低下検出時には、パケット送信量を直ちには削減せずにそのまま保持し、

その後のスループット低下がさらに悪化した場合にのみパケット送信量を削減して通信帯域を削減することを特徴とする帯域制御方法。

【請求項６】　請求項５記載の帯域制御方法において、

ユーザレベルあるいはアプリケーション種別に応じて帯域を優先的に割当てる優先通信が必要か否か判断し、

優先通信が不要な場合には、スループット低下検出時にパケット送信量を直ちに削減し、

優先通信が必要な場合には、スループット低下検出時にパケット送信量を直ちには削減せずにそのまま保持し、

その後のスループット低下がさらに悪化した場合にのみ

2

パケット送信量を削減して通信帯域を削減することを特徴とする帯域制御方法。

【請求項７】　請求項５または６記載の帯域制御方法において、

任意の第１の測定区間でのスループットと、その直後の第２の測定区間でのスループットとの差から第１のスループット変化量を算出し、

算出された第１のスループット変化量が所定しきい値以上の場合に、スループットが低下していると判断し、

第２の測定区間でのスループットと、その直後の第３の測定区間でのスループットとの差から第２のスループット変化量を算出し、

算出された第１および第２のスループット変化量が所定しきい値以上の場合には、スループット低下がさらに悪化していると判断することを特徴とする帯域制御方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、帯域制御方法に関し、特に通信プロトコル処理を実行するノードにおけるプロトコルレイヤ４の帯域を制御する場合の帯域制御方法に関するものである。

【０００２】

【従来の技術】現在、インターネット／イントラネットではベストエフォート型サービスが主要なサービス形態であり、各コネクションには時々刻々変化するネットワーク状況に応じた帯域が提供される。代表的通信プロトコルであるＴＣＰでは、パケット紛失事象の検出などを基にしてネットワーク状況の変化を検出し、送信量を調整する。

【０００３】このために、相手ノードの受信能力を示すパラメータである受信告知ウィンドウ値（ａｄｗｎｄ）の他に、ネットワークの転送能力の推定値を示すパラメータである輻輳ウィンドウ値（ｃｗｎｄ）を用いて、ｍｉｎ（ａｄｗｎｄ，ｃｗｎｄ）の範囲で送信する。ただし、ｍｉｎ（Ａ，Ｂ）はＡとＢのいずれか小さい方の値を採ることを示している。

【０００４】送信量の調整のための帯域制御はｃｗｎｄを中心として実現され、前述のパラメータの他にスロースタート閾値（ｓｓｔｈｒｅｓｈ）がある。これらのパラメータに基づいてパケット送信が行われるが、スロースタートフェーズと輻輳回避フェーズと呼ばれる２つのフェーズがあり、制御の方法が異なる。

【０００５】その方式の概要は次のようなものである。なお、帯域制御の各パラメータの単位は、ＴＣＰの実装ではバイト数を用いているが、以下では説明を容易にするためパケット数で表現する。まず、コネクション設定後、帯域制御のパラメータを、ｓｓｔｈｒｅｓｈ＝ａｄｗｎｄ、ｃｗｎｄ＝１に初期設定する。

【０００６】そしてスロースタートフェーズに入り、１個のデータパケット（ＤＴパケット）の送信を行い、確

認応答パケット（ＡＣＫパケット）の受信を待つ。一定時間内にＡＣＫパケットを受信したら、ｃｗｎｄを＋１し、次には２パケットの送信を行い、以降、ＡＣＫパケットを受信するごとに確認したＤＴパケット数だけ、次の送信可能量であるｃｗｎｄがｓｓｔｈｒｅｓｈになるまで増加させる。

【０００７】したがって、送信可能量ｃｗｎｄは、１，２，４，８，…と言った具合に増加していく。ｃｗｎｄがｓｓｔｈｒｅｓｈに達したら、輻輳回避フェーズに入る。輻輳回避フェーズでは、ＡＣＫパケット受信ごとにｃｗｎｄを１／ｃｗｎｄだけ増加させるので、スロースタートフェーズと比較して、遥かに緩やかな増加となる。

【０００８】代表的なＴＣＰの実装では、ＤＴパケット紛失事象は、確認応答が受信されず時間監視がタイムアウトした時、または一定数（通常は３）以上の重複ＡＣＫを受信した時に、パケット紛失が発生したものと判断する。なお、重複ＡＣＫ受信とは、同一の受信シーケンス番号を有するＡＣＫパケットを連続して複数個受信することをいう。

【０００９】ＤＴパケット紛失と判断した時は、帯域制御のパラメータを調整し、ｓｓｔｈｒｅｓｈはｍｉｎ（ｃｗｎｄ，ａｄｗｎｄ）／２とする。またｃｗｎｄについて、重複ＡＣＫ検出の場合には、新ｃｗｎｄ＝旧ｃｗｎｄ／２とし、タイムアウトの場合には、新ｃｗｎｄ＝１とするものとなっている（例えば、W.R.Stevens著、TCP Illustrated Vol.1 のChapter 21、Addison Wesley、1994など参照）

【００１０】一方、誤り回復の迅速化を狙いとした選択的応答（ＳＡＣＫ）の機能がＴＣＰのオプションとして１９９６年にインターネット・ドキュメントであるＲＦＣ２０１８で規定された。ＳＡＣＫ情報は、受信側でパケット紛失により非連続受信となったパケット全てについて、受信できた区間をＡＣＫパケットに付加して具体的に通知するものである。送信側では、この情報を使用して、紛失パケットのみを再送する。

【００１１】帯域制御（輻輳制御）との関係については、ＲＦＣ２０１８では既存機能が保存されるべきとあるので、ＳＡＣＫ情報によりパケット紛失を検出した場合には、前述の重複ＡＣＫ検出時と同様のパラメータ調整を行う。以上のように、重複ＡＣＫ、応答確認タイムアウト、ＳＡＣＫ情報によりパケット紛失事象を検出した場合、そのコネクションを使用しているユーザレベルやアプリケーション種別によらず、一律に送信量を削減する制御を行うものとなっていた。

【００１２】

【発明が解決しようとする課題】しかしながら、このような従来の帯域制御方法、特に、現状の各種ホスト／端末装置で実現されているＴＣＰでは、パケット紛失事象が発生するようなネットワーク混雑時には、アプリケーションやユーザの通信要求条件に依らず、各コネクションの通信帯域を一律に削減するベストエフォート型サービスを提供するものとなっているため、インターネットの混雑時には、例えば、重要／緊急などの優先度の高い通信のコネクションで帯域を確保することができないという問題点があった。

【００１３】すなわち、情報通信サービスの多様化・高度化に伴い、ベストエフォート型サービスにおいて、いわゆるＱｏＳ（サービス品質）の要求条件の一つとして、通信帯域の利用についての優先度付けを行い、きめ細かいサービス実現が必要となりつつある。例えば、社内利用を想定したイントラネットにおいて、業務種別や職位に応じて通信帯域の割当の優先度をつけて、重要通信や緊急通信でのスループット確保や応答時間短縮を図ることが考えられるが、現状のＴＣＰの帯域制御では実現できない。

【００１４】本発明はこのような課題を解決するためのものであり、通信帯域の利用についての優先度付けに基づいたきめ細かいサービスを実現できる帯域制御方法を提供することを目的としている。

【００１５】

【課題を解決するための手段】このような目的を達成するために、本発明による帯域制御方法のうち、請求項１の発明は、相手ノードの受信能力を示す受信告知ウィンドウ値と、ネットワークの転送能力を推定した輻輳ウィンドウ値とのいずれか小さい値に基づいて、パケット送信に用いる通信帯域を設定する通信プロトコルにおいて、パケット紛失事象検出時には、輻輳ウィンドウ値を直ちには削減せずにそのまま保持し、その後のパケット紛失状況が悪化した場合にのみ輻輳ウィンドウ値を削減して通信帯域を削減するようにしたものである。したがって、パケット紛失事象検出時には、直ちに通信帯域が削減されず現状のまま保持され、その後のパケット紛失状況が悪化した場合にのみ通信帯域が削減される。

【００１６】また、請求項２の発明は、請求項１の発明において、ユーザレベルあるいはアプリケーション種別に応じて帯域を優先的に割当てる優先通信が必要か否か判断し、優先通信が不要な場合には、パケット紛失事象検出時に輻輳ウィンドウ値を直ちに削減し、優先通信が必要な場合には、パケット紛失事象検出時に輻輳ウィンドウ値を直ちには削減せずにそのまま保持し、その後のパケット紛失状況が悪化した場合にのみ輻輳ウィンドウ値を削減して通信帯域を削減するようにしたものである。したがって、パケット紛失事象検出時、優先通信が不要なコネクションについては直ちに通信帯域が削減され、これにより解放された通信帯域を用いて優先通信が必要なコネクションの通信帯域が現状のまま保持され、その後のパケット紛失状況が悪化した場合になって初めて、優先通信が必要なコネクションの通信帯域が削減される。

【００１７】また、請求項３の発明は、請求項１または２の発明において、所定の監視区間ごとにパケット紛失量を算出し、任意の監視区間での第１のパケット紛失量と、その直後の監視区間での第２のパケット紛失量とを比較し、第２のパケット紛失量が第１のパケット紛失量より大きい場合には、パケット紛失状況が悪化していると判断するようにしたものである。また、請求項４の発明は、請求項３の発明において、パケット紛失事象検出時点から送信済みのパケットすべてについて受信側から送達確認通知された時点までの区間を監視区間とするようにしたものである。

【００１８】また、請求項５の発明は、ネットワークの負荷状況を示すスループットを所定長の測定区間ごとに測定し、その変化に応じてパケット送信量を調整する通信プロトコルにおいて、スループットの低下検出時には、パケット送信量を直ちには削減せずにそのまま保持し、その後のスループット低下がさらに悪化した場合にのみパケット送信量を削減して通信帯域を削減するようにしたものである。したがって、スループット低下検出時には、直ちに通信帯域が削減されず現状のまま保持され、その後のスループット低下がさらに悪化した場合にのみ通信帯域が削減される。

【００１９】また、請求項６の発明は、請求項５の発明において、ユーザレベルあるいはアプリケーション種別に応じて帯域を優先的に割当てる優先通信が必要か否か判断し、優先通信が不要な場合には、スループット低下検出時にパケット送信量を直ちに削減し、優先通信が必要な場合には、スループット低下検出時にパケット送信量を直ちには削減せずにそのまま保持し、その後のスループット低下がさらに悪化した場合にのみパケット送信量を削減して通信帯域を削減するようにしたものである。したがって、スループット低下検出時、優先通信が不要なコネクションについては直ちに通信帯域が削減され、これにより生じた通信帯域を用いて優先通信が必要なコネクションの通信帯域が現状のまま保持され、その後のスループット低下がさらに悪化した場合になって初めて、優先通信が必要なコネクションの通信帯域が削減される。

【００２０】また、請求項７の発明は、請求項５または６の発明において、任意の第１の測定区間でのスループットと、その直後の第２の測定区間でのスループットとの差から第１のスループット変化量を算出し、算出された第１のスループット変化量が所定しきい値以上の場合に、スループットが低下していると判断し、第２の測定区間でのスループットと、その直後の第３の測定区間でのスループットとの差から第２のスループット変化量を算出し、算出された第１および第２のスループット変化量が所定しきい値以上の場合には、スループット低下がさらに悪化していると判断するようにしたものである。

【００２１】

【発明の実施の形態】次に、本発明について図面を参照して説明する。図１は本発明の第１の実施の形態である通信システムのブロック図である。同図において、エンドノード１は、情報通信網２に接続され、相手のエンドノード１と通信するノード、情報通信網２は通信回線および中継ノードから構成されるネットワークである。

【００２２】エンドノード１は、下位レイヤ処理部１１、レイヤ４処理部１２、アプリケーション処理部１３から構成されている。下位レイヤ処理部１１は、プロトコルレイヤ３以下の処理、すなわち、通信回線との電気的整合などのレイヤ１、フレームの組立／分解などのレイヤ２、およびルーティングなどのレイヤ３（ここではＩＰとする）の各処理を行う。

【００２３】レイヤ４処理部１２は、レイヤ４（ここではＴＣＰとする）のコネクションの設定解放やフロー制御などに基づいたデータ送受信の処理を行う。この場合、プロトコルレイヤ４（ＴＣＰ）用の帯域制御のパラメータとして、受信告知ウィンドウ値（ａｄｗｎｄ）、輻輳ウィンドウ値（ｃｗｎｄ）、スロースタート閾値（ｓｓｔｈｒｅｓｈ）の他に、パケット紛失数をカウントする機能、パケット紛失数を保持する変数ｌｏｓｓ１とｌｏｓｓ２、パケット紛失事象検出時の送信済みで未確認の最大送信シーケンス番号を保持する変数ｈｉ＿ｃｈｋを用いる。

【００２４】送信可能なパケット数は、ｍｉｎ（ｃｗｎｄ，ａｄｗｎｄ）であり、各パラメータの変動に伴い通信中に変化していく。ただし、ｍｉｎ（Ａ，Ｂ）はＡとＢの小さい方の値を採ることを示す。また、上位レイヤがＯＮ／ＯＦＦを設定し、ＴＣＰでの帯域制御の方法を選択するための優先通信フラグを持つ。

【００２５】本発明の第１の実施の形態では、優先通信フラグがＯＮのコネクションにおいてパケット紛失事象が検出された場合には、直ちに帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄを削減せずに保持し、その後の紛失状況が悪化した場合にのみ帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄを削減するものである。また、パケット紛失事象が検出された時点での送信済みでかつ送達未確認のパケットすべてについて送達確認されるまでを監視区間とし、隣接する２つの監視区間に紛失したそれぞれの紛失パケット総数を比較することにより、パケット紛失事象検出後の紛失状況が悪化しているか否かを判断するものである。

【００２６】次に、図２～４を参照して、本発明の第１の実施の形態による動作として、帯域制御を中心としたデータ送信の動作について説明する。図２は通信シーケンス全体の概略を示すシーケンス説明図、図３は帯域制御全体の概略を示すフローチャート、図４は優先通信時の帯域制御処理を示す状態遷移表である。以下では、クライアント側（受信側）からの要求に応じて、サーバ側（送信側）からパケット転送を行う場合を例に説明す

る。

【００２７】まず、クライアント側のレイヤ４処理部１２（ＴＣＰ）は、アプリケーション処理部１３（上位ＡＰ）からの通信開始要求に基づいて、コネクション設定要求およびコネクション設定応答を、サーバ側のレイヤ４処理部１２（ＴＣＰ）とやり取りすることにより、コネクションの設定を行う。このとき、ＴＣＰコネクションでＳＡＣＫオプションを使用したい場合には、コネクション設定要求パケットにＳＡＣＫ表示を付加して、送信側と受信側でネゴシエーションしておく。

【００２８】なお、クライアント側のレイヤ４処理部１２（ＴＣＰ）は、アプリケーション処理部１３（上位ＡＰ）からの通信終了要求に基づいて、コネクション開放要求およびコネクション開放応答を、サーバ側のレイヤ４処理部１２（ＴＣＰ）とやり取りすることにより、コネクションの開放を行う。

【００２９】図２に示すように、コネクション設定完了後、サーバ側はデータ転送可能な通常（Ｎ）状態となる（図３：ステップ３１）。サーバ側の上位ＡＰは、クライアント側の上位ＡＰから通知されるユーザＩＤまたは要求アプリケーションＩＤに基づいて、設定したＴＣＰコネクションで優先的な帯域制御を行うか否か判断する。

【００３０】ここで、通知されたＩＤが、優先通信の必要なコネクションを示す場合には、優先通信フラグをＯＮに設定し、優先通信の必要でないコネクションを示す場合には、優先通信フラグをＯＦＦに設定しておく。したがって、以降のデータ転送の通常（Ｎ）状態において、サーバ側のレイヤ４処理部１２により、ＴＣＰのＤＴパケットの紛失事象が検出された場合には、この優先通信フラグのＯＮ／ＯＦＦにより、そのコネクションに優先通信が必要か否か判断される。

【００３１】図３に示すように、データ転送の通常（Ｎ）状態（ステップ３１）において、サーバ側のレイヤ４処理部１２により、ＴＣＰのＤＴパケットの紛失事象が検出された場合、その優先通信フラグがチェックされる（ステップ３２）。ここで、優先通信フラグがＯＦＦならば、従来と同様の通常帯域制御が指定されているコネクションであると判断して、直ちに帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄの削減を行う（ステップ３３）。

【００３２】一方、優先通信フラグがＯＮの場合には、優先通信が必要なコネクションであると判断して、以下のような帯域制御処理を実行する。なお、パケット紛失事象は、重複ＡＣＫ、ＳＡＣＫ（選択的応答）情報、送達応答確認タイムアウトにより把握できる。また、紛失パケット数Ｌは、重複ＡＣＫ、送達確認タイムアウトの場合は１となるが、ＳＡＣＫ情報の場合はその情報から判明する新たな紛失数Ｌとなる。

【００３３】通常（Ｎ）状態で、重複ＡＣＫ受信時またはＳＡＣＫ受信時には、監視区間が開始され、その時点での送信済みで未確認の最大送信シーケンス番号を変数ｈｉ＿ｃｈｋへ設定し、変数ｌｏｓｓ１へ紛失パケット数Ｌを設定し、ｃｗｎｄとｓｓｔｈｒｅｓｈの値を削減することなく、監視１（Ｍ１）状態（ステップ３４）へ遷移する。またＮ状態において、送達確認タイムアウト時にも監視区間が開始され、その時点での送信済みで未確認の最大送信シーケンス番号を変数ｈｉ＿ｃｈｋへ設定し、変数ｌｏｓｓ１へ１を設定し、ｃｗｎｄとｓｓｔｈｒｅｓｈの値を削減することなく、Ｍ１状態（ステップ３４）へ遷移する。

【００３４】Ｍ１状態において、ＡＣＫ受信時の処理は次のようになる。受信したＡＣＫパケットの受信シーケンス番号（ＡＣＫｓｅｑ）が、ｈｉ＿ｃｈｋ＞ＡＣＫｓｅｑの場合、すなわち送信済みで未確認のパケットすべてについてまだ送達確認されていない場合には、ｌｏｓｓ１の変更はなく、Ｍ１状態のままとし、監視区間を継続する。

【００３５】一方、ｈｉ＿ｃｈｋ≦ＡＣＫｓｅｑの場合、すなわち送信済みで未確認のパケットすべてについて送達確認された場合には、その時点での送信済みで未確認の最大送信シーケンス番号をｈｉ＿ｃｈｋへ設定し、変数ｌｏｓｓ２を０とし、監視２（Ｍ２）状態（ステップ３５）へ遷移する。これにより、最初の監視区間が終了して、次の監視区間が開始される。

【００３６】また、Ｍ１状態において、重複ＡＣＫ受信時またはＳＡＣＫ受信時の処理は次のようになる。受信するＡＣＫパケットの受信シーケンス番号（ＡＣＫｓｅｑ）が、ｈｉ＿ｃｈｋ＞ＡＣＫｓｅｑの場合、すなわち送信済みで未確認のパケットすべてについてまだ送達確認されていない場合には、新たなパケット紛失の検出があればｌｏｓｓ１へ紛失数Ｌを加算し、Ｍ１状態のままとし、監視区間を継続する。

【００３７】一方、ｈｉ＿ｃｈｋ≦ＡＣＫｓｅｑの場合、すなわち送信済みで未確認のパケットすべてについて送達確認された場合には、その時点での送信済みで未確認の最大送信シーケンス番号をｈｉ＿ｃｈｋへ設定し、新たなパケット紛失の検出があればｌｏｓｓ２へ紛失数Ｌを設定して、Ｍ２状態へ遷移する。これにより、最初の監視区間が終了して、次の監視区間が開始される。さらに、Ｍ１状態で、送達確認タイムアウト時は、ｌｏｓｓ１を＋１し、Ｍ１状態のままとし、監視区間を継続する。

【００３８】また、Ｍ２状態において、ＡＣＫ受信時の処理は次のようになる。受信するＡＣＫパケットの受信シーケンス番号（ＡＣＫｓｅｑ）が、ｈｉ＿ｃｈｋ＞ＡＣＫｓｅｑの場合、すなわち送信済みで未確認のパケットすべてについてまだ送達確認されていない場合には、ｌｏｓｓ２の変更はなく、Ｍ２状態のままとし、監視区間を継続する。

【００３９】一方、ｈｉ＿ｃｈｋ≦ＡＣＫｓｅｑの場合、すなわち送信済みで未確認のパケットすべてについて送達確認された場合には、前後２つの監視区間が終了したと判断して、ｌｏｓｓ１とｌｏｓｓ２の値を比較する（ステップ３６）。ここで、ｌｏｓｓ２＝０ならば、後続の監視区間において紛失パケットが発生していないことから、輻輳状況は解消したと判断して、帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄは削減せず、通常（Ｎ）状態（ステップ３１）へ遷移する。

【００４０】また、ｌｏｓｓ１≧ｌｏｓｓ２ならば、後続の監視区間での紛失パケット総数が低減していることから、輻輳状況は改善しつつあると考えられるため、帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄは削減せず、さらに様子を見るため、新たな監視区間を開始する。このため、ｌｏｓｓ２の値をｌｏｓｓ１へ設定し、ｌｏｓｓ２＝０とし、その時点での送信済みで未確認の最大送信シーケンス番号をｈｉ＿ｃｈｋへ設定して、Ｍ２状態（ステップ３５）へ戻り、監視区間を継続する。

【００４１】また、ｌｏｓｓ１＜ｌｏｓｓ２ならば、後続の監視区間での紛失パケット総数が増加していることから、ネットワークの輻輳状況が悪化していると考えられるため、帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄを削減（ステップ３７）した後に、Ｎ状態（ステップ３１）へ遷移する。

【００４２】また、Ｍ２状態において、重複ＡＣＫ受信時またはＳＡＣＫ受信時の処理は次のようになる。受信するＡＣＫパケットの受信シーケンス番号（ＡＣＫｓｅｑ）が、ｈｉ＿ｃｈｋ＞ＡＣＫｓｅｑの場合、すなわち送信済みで未確認のパケットすべてについてまだ送達確認されていない場合には、新たなパケット紛失の検出があればｌｏｓｓ２へ紛失数Ｌを加算し、Ｍ２状態のままとし、監視区間を継続する。

【００４３】ｈｉ＿ｃｈｋ≦ＡＣＫｓｅｑの場合、すなわち送信済みで未確認のパケットすべてについて送達確認された場合には、前後２つの監視区間が終了したと判断して、ｌｏｓｓ１とｌｏｓｓ２の値を比較する（ステップ３８）。ここで、ｌｏｓｓ１≧ｌｏｓｓ２ならば、後続の監視区間での紛失パケット総数が低減していることから、輻輳状況は改善しつつあると考えられるため、帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄは削減しない。

【００４４】そして、さらに様子を見るため、ｌｏｓｓ２の値をｌｏｓｓ１へ設定するとともに、その時点での送信済みで未確認の最大送信シーケンス番号をｈｉ＿ｃｈｋへ設定し、新たなパケット紛失の検出があればｌｏｓｓ２へ紛失数Ｌを設定して、Ｍ２状態（ステップ３５）へ戻り、新たな監視区間を開始する。

【００４５】また、ｌｏｓｓ１＜ｌｏｓｓ２ならば、後続の監視区間での紛失パケット総数が増加していることから、ネットワークの輻輳状況が悪化していると考えられるため、帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄを削減（ステップ３９）した後に、Ｎ状態（ステップ３１）へ遷移する。

【００４６】さらに、Ｍ２状態で、送達確認タイムアウト時は、ｌｏｓｓ２を＋１し、Ｍ２状態のままとし、監視区間を継続する。なお、ステップ３６，３８での帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄの削減は、ここでは従来方式と同様に、ｓｓｔｈｒｅｓｈ＝ｍｉｎ（ｃｗｎｄ，ａｄｗｎｄ）／２、新ｃｗｎｄ＝旧ｃｗｎｄ／２とする。

【００４７】このように、本発明の第１の実施の形態では、パケット紛失事象が検出された場合には、直ちに帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄを削減せずに保持し、その後の紛失状況が悪化した場合にのみ帯域制御パラメータｓｓｔｈｒｅｓｈとｃｗｎｄを削減するようにした。したがって、図５（ａ）に示す従来の帯域制御方法のように、パケット紛失事象の検出に応じて一律に帯域削減を行う場合と比較して、図５（ｂ）に示すように、重要／緊急などの優先度の高い通信のコネクションで帯域を確保することができ、通信帯域の利用についての優先度付けに基づいたきめ細かいサービスを実現できる。

【００４８】また、パケット紛失事象が検出された時点での送信済みでかつ送達未確認のパケットすべてについて送達確認されるまでを監視区間とし、任意の監視区間とその直後の監視区間で紛失したそれぞれの紛失パケット総数ｌｏｓｓ１，ｌｏｓｓ２を比較するようにしたので、パケット紛失事象検出後の紛失状況が悪化しているか否かを正確に判断できる。

【００４９】また、受信側（クライアント側）から通知されたユーザＩＤまたは要求アプリケーションＩＤに基づき、送信側（サーバ側）でそのコネクションへの優先的な帯域制御を行うか否か判断するようにしたので、多数の受信側に対する無制限な優先通信を容易に抑制することができるとともに、優先通信を行わない従来と同様のコネクションについてはパケット紛失事象の検出に応じて直ちに帯域削減が行われ、優先通信を行うための帯域を確保できる。

【００５０】次に、本発明の第２の実施の形態について説明する。ここでは、ネットワークの輻輳状況をパケット紛失事象でなく、スループットの変化状況から把握して、送信量を調整する通信プロトコル（例えば、ＴＣＰ－Ｖｅｇａｓと呼ばれるＴＣＰバージョン）の場合について説明する。なお、考え方は前述と同様であり、ここでは簡単に説明する。

【００５１】まず、所定長の測定区間ごとにスループットを算出する。スループットを算出する場合、まず、任意の測定区間でのパケット送信開始から、その測定区間に送信した全送信パケットに対して受信側から送達確認通知されるまでの所要時間を計時する。そして、その測

定区間に送信したパケット送信量を所要時間で除算することによりスループットを算出する。

【００５２】続いて、任意の測定区間で算出したスループットをＰ１とするとともに、その直後の測定区間で算出したスループットをＰ２とし、両測定区間におけるスループット変化量Ｄを、両者の差すなわちＤ＝Ｐ１－Ｐ２で求める。この場合、所定のしきい値ａを定めて、スループットが低下しているか否かの状況判断に使用する。

【００５３】ここで、優先通信フラグがＯＮの場合の動作は次のようになる。Ｄ＜ａならば、スループットが良好であると判断して、パケット送信量を増加させる。また、Ｄ≧ａならば、スループットが低下していると判断して、現状の送信量の値を保持して次期間の様子を見る。

【００５４】続いて、このＰ２の値をＰ１として保存しておき、次期間のスループットをＰ２として求める。そして、再びスループット変化量Ｄを算出してしきい値ａとの比較を行い、再びＤ≧ａならば、スループット低下がさらに悪化していると判断して、現状のパケット送信量の値を削減する。

【００５５】一方、優先通信フラグがＯＦＦの場合の動作は次のようになる。前述と同様にして算出したスループット変化量Ｄとしきい値ａとを比較し、Ｄ＜ａならば、スループットが良好であると判断して、送信量を増加させる。また、Ｄ≧ａならば、スループットが低下して、ネットワークの負荷が高くなりつつあるものとして、直ちに現状の送信量の値を削減する。

【００５６】したがって、スループットの変化状況から把握して、送信量を調整する通信プロトコルでは、スループット低下が検出された場合には、直ちにパケット送信量を削減せずに保持し、その後のスループット低下がさらに悪化した場合にのみパケット送信量を削減するようにしたので、従来のようにパケット紛失事象の検出に応じて一律に帯域削減を行う場合と比較して、重要／緊急などの優先度の高い通信のコネクションで帯域を確保することができ、通信帯域の利用についての優先度付けに基づいたきめ細かいサービスを実現できる。

【００５７】

【発明の効果】以上説明したように、本発明では、ネットワークの輻輳状況が検出された場合には、優先通信を行うコネクションにおいて、直ちにパケット送信量を削減せずそのまま保持し、その後の輻輳状況がさらに悪化した場合にのみ優先通信を行うコネクションの通信帯域を削減するようにしたものである。したがって、従来のようにパケット紛失事象の検出に応じて一律に帯域削減を行う場合と比較して、重要／緊急などの優先度の高い通信のコネクションで帯域を確保することができるとともに、さらには、ベストエフォート型サービスにおけるＱｏＳ（サービス品質）の多様化に対応でき、通信帯域の利用についての優先度付けに基づいたきめ細かいサービスを実現できる。例えば、社内利用を想定したイントラネットにおいて、業務種別やユーザ種別に応じて通信帯域の割当の優先度をつけて、重要通信や緊急通信でのスループット確保や応答時間短縮を図ることができ、アグレッシブなベストエフォート型サービスを実現できる。

【００５８】また、ユーザレベルあるいはアプリケーション種別に応じて帯域を優先的に割当てる優先通信が必要か否か判断し、優先通信が不要なコネクションについては、輻輳状況検出時にパケット送信量を直ちに削減し、優先通信が必要なコネクションについては、輻輳状況検出時にパケット送信量を直ちには削減せずにそのまま保持し、その後の輻輳状況が悪化した場合にのみパケット送信量を削減して通信帯域を削減するようにしたものである。したがって、ネットワークで輻輳状況が発生し、使用可能な通信帯域が低減した場合でも、輻輳発生に応じて直ちに削減された優先通信が不要なコネクションの通信帯域を用いて、優先通信が必要なコネクションの通信帯域が現状のまま保持される。さらに、ユーザレベルあるいはアプリケーション種別に応じて帯域を優先的に割当てる優先通信が必要か否かを送信側で判断でき、多数の受信側に対する無制限な優先通信を容易に抑制することができるとともに、優先通信を行わない従来と同様のコネクションについてはパケット紛失事象の検出に応じて直ちに帯域削減が行われ、優先通信を行うための帯域を確保できる。

【００５９】また、所定の監視区間ごとにパケット紛失量を算出し、任意の監視区間での第１のパケット紛失量と、その直後の監視区間での第２のパケット紛失量とを比較し、第２のパケット紛失量が第１のパケット紛失量より大きい場合には、パケット紛失状況が悪化していると判断するようにし、さらには、パケット紛失事象検出時点から送信済みのパケットすべてについて受信側から送達確認通知された時点までの区間を監視区間とするようにしたので、ネットワークの輻輳状況をパケット紛失事象により検出し、そのパケット送信量を調整する通信プロトコルにおいて、パケット紛失事象検出後の紛失状況が悪化しているか否かを正確に判断できる。

【００６０】また、任意の第１の測定区間でのスループットと、その直後の第２の測定区間でのスループットとの差から第１のスループット変化量を算出し、算出された第１のスループット変化量が所定しきい値以上の場合に、スループットが低下していると判断し、第２の測定区間でのスループットと、その直後の第３の測定区間でのスループットとの差から第２のスループット変化量を算出し、算出された第１および第２のスループット変化量が所定しきい値以上の場合には、スループット低下がさらに悪化していると判断するようにしたので、ネットワークの輻輳状況をスループットの変化により検出し、

13

そのパケット送信量を調整する通信プロトコルにおいて、スループットの低下、およびその後のさらなる悪化を正確に判断できる。

【図面の簡単な説明】

【図１】　本発明の一実施の形態であるデータ通信システムのブロック図である。

【図２】　通信シーケンス全体の概略を示すシーケンス説明図である。

【図３】　帯域制御全体の概略を示すフローチャートである。

14

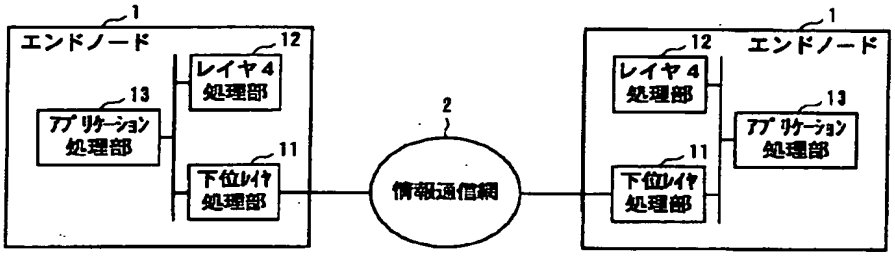
【図４】　優先通信時の帯域制御処理を示す状態遷移表である。

【図５】　本発明の動作概念を示す説明図である。

【符号の説明】

１…エンドノード、２…情報通信網、１１…下位レイヤ処理部、１２…レイヤ４処理部、１３…アプリケーション処理部。

【図１】

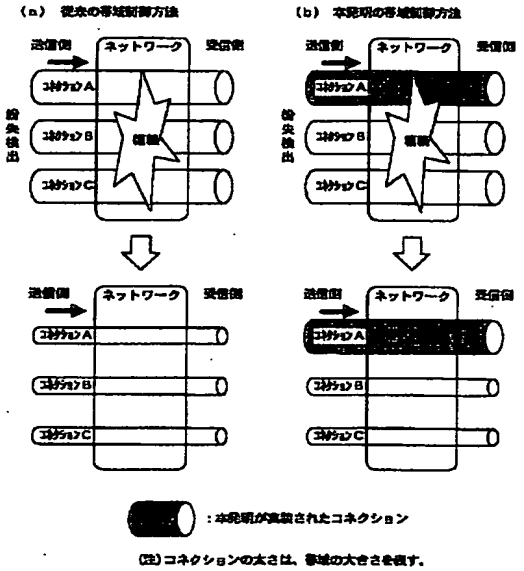


【図４】

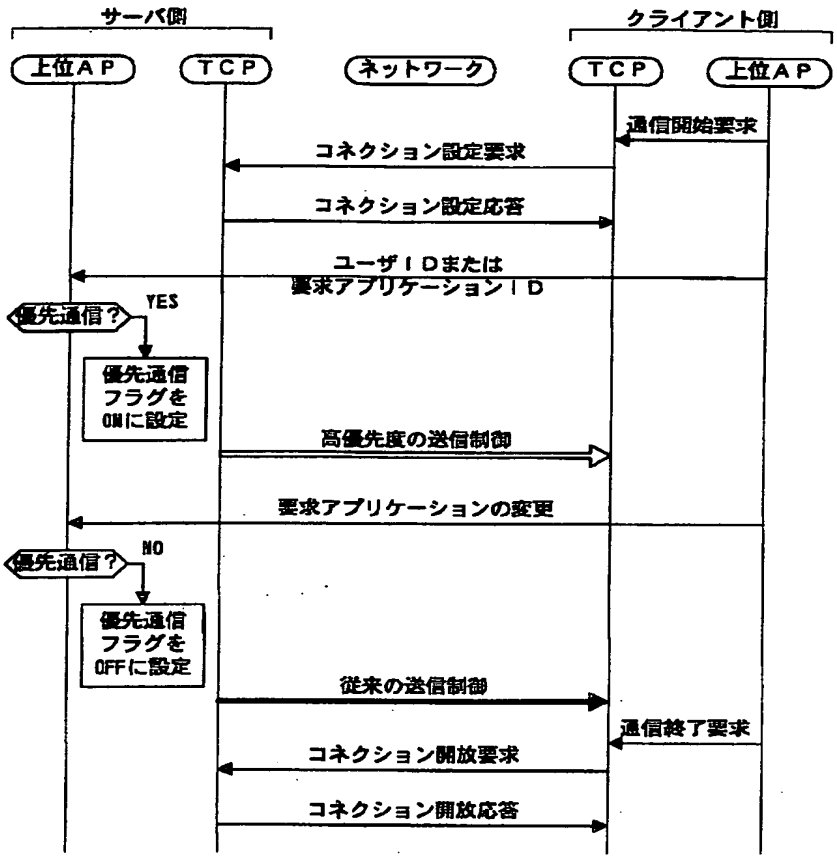
状態 / 遷移規則	通常（N）	監視１（M１）	監視２（M２）
ACK受信	・従来動作 ［スロースタート，輻輳回避フェーズの通信動作］ ・状態Nのまま	(1) bl_ctk＞ACKseqの場合 ・状態M１のまま (2) bl_ctk≦ACKseqの場合 ・bl_ctk設定 ・loss1←0 ・状態M２へ	(1) bl_ctk＞ACKseqの場合 ・状態M２のまま (2) bl_ctk≦ACKseqの場合 (a) loss2＝0の場合 ・状態Nへ (b) loss1≧loss2 ・loss1←loss2 ・bl_ctk設定 ・loss2←0 ・状態M２のまま (c) loss1＜loss2の場合 ・ssthresh，cwndを調整 ・状態Nへ
重複ACK受信 または SACK受信	・bl_ctk設定 ・loss1←L ・状態M１へ	(1) bl_ctk＞ACKseqの場合 ・loss1←loss1+L ・状態M１のまま (2) bl_ctk≦ACKseqの場合 ・bl_ctk設定 ・loss2←L ・状態M２へ	(1) bl_ctk＞ACKseqの場合 ・loss2←loss2+L ・状態M２のまま (2) bl_ctk≦ACKseqの場合 (a) loss1≧loss2の場合 ・loss1←loss2 ・bl_ctk設定 ・loss2←L ・状態M２のまま (c) loss1＜loss2の場合 ・ssthresh，cwndを調整 ・状態Nへ
送信確認タイムアウト	・bl_ctk設定 ・loss1←1 ・状態M１へ	・loss1←loss1+1 ・状態M１のまま	・loss2←loss2+1 ・状態M２のまま

【記号】
bl_ctk設定：送信済みで未確認の最大送信シーケンス番号を，変数bl_ctkへ設定すること。
L：仮定した損失パケット数
（重複ACKの場合は1，SACKの場合は新規に判明した損失数）

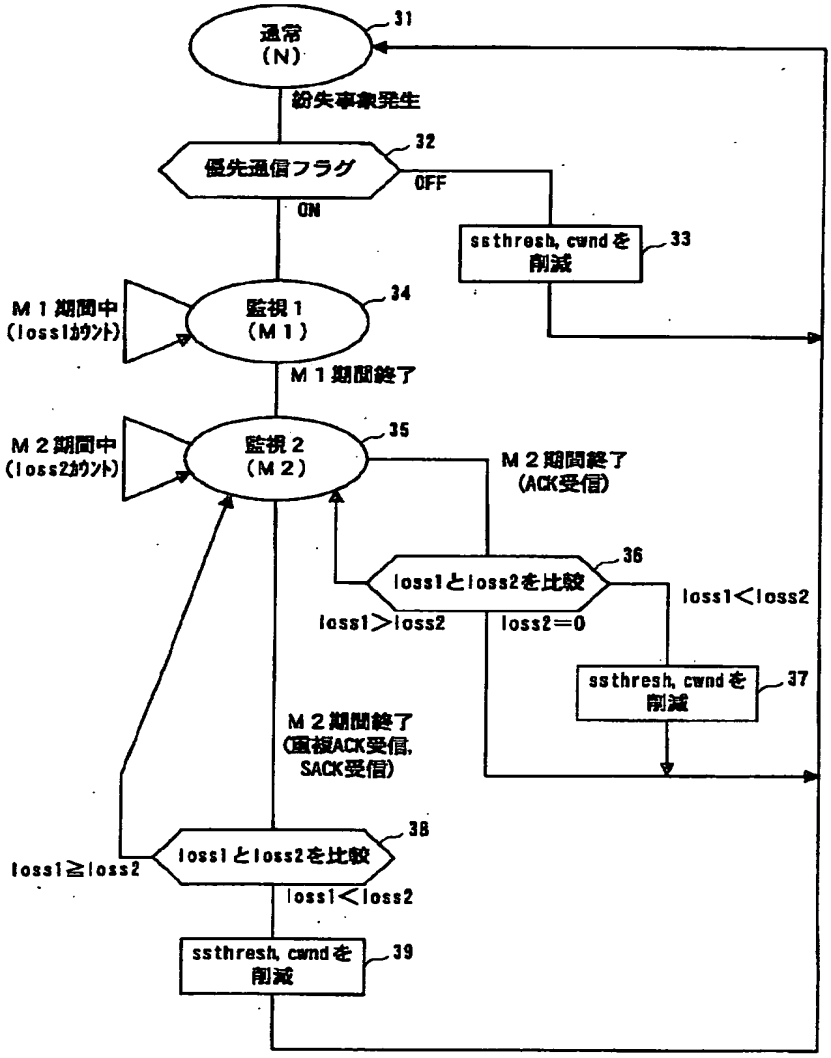
【図５】



【図２】


サーバ側
クライアント側
上位ＡＰ
ＴＣＰ
ネットワーク
ＴＣＰ
上位ＡＰ
通信開始要求
コネクション設定要求
コネクション設定応答
ユーザＩＤまたは
要求アプリケーションＩＤ
優先通信？
YES
優先通信
フラグを
ONに設定
高優先度の送信制御
要求アプリケーションの変更
優先通信？
NO
優先通信
フラグを
OFFに設定
従来の送信制御
通信終了要求
コネクション開放要求
コネクション開放応答

【図３】


通常
(N)
31
紛失事象発生
優先通信フラグ
32
OFF
ON
ssthresh, cwndを
削減
33
M１期間中
(loss1カウント)
監視１
(M１)
34
M１期間終了
M２期間中
(loss2カウント)
監視２
(M２)
35
M２期間終了
(ACK受信)
loss1とloss2を比較
36
loss1＜loss2
loss1＞loss2
loss2＝0
ssthresh, cwndを
削減
37
M２期間終了
(重複ACK受信,
SACK受信)
loss1とloss2を比較
38
loss1≧loss2
loss1＜loss2
ssthresh, cwndを
削減
39